



ELSEVIER

SCIENCE @ DIRECT®

Applied Numerical Mathematics 53 (2005) 391–408

APPLIED
NUMERICAL
MATHEMATICSwww.elsevier.com/locate/apnum

Numerical analysis of structure preserving Nyström methods for Hamiltonian systems

Andreas Pfeiffer^a, Martin Arnold^{b,*}^a *German Aerospace Center (DLR) Oberpfaffenhofen, Institute of Robotics and Mechatronics,
P.O. Box 1116, 82230 Wessling, Germany*^b *Martin Luther University Halle-Wittenberg, Department of Mathematics and Computer Science,
Institute of Numerical Mathematics, 06099 Halle (Saale), Germany*

Abstract

Energy conservation is an essential property of conservative mechanical systems that should be carried over to the numerical solution. Betsch and Steinmann proposed recently perturbed potentials to achieve energy conservation in the time integration of N -body problems by Galerkin methods. In the present paper this approach is generalised to Nyström methods for Hamiltonian systems. A detailed analysis shows that energy conservation by perturbed potential functions does not affect the feasibility and (high) order of convergence of Nyström methods. Symmetry and reversibility properties are left unchanged as well. The theoretical results are illustrated by numerical tests indicating clearly the benefits of energy conserving methods in long-term simulations.

© 2004 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: Geometric integration; Hamiltonian system; Nyström methods; Galerkin methods; Collocation; Energy conservation

1. Introduction

The time integration of ordinary differential equations is a classical topic of numerical mathematics. In the traditional approach one-step methods like Runge–Kutta methods and multistep methods like Adams methods or BDF are constructed that combine a high order of convergence with small numerical effort per time step and favourable stability properties for linear problems [9].

* Corresponding author.

E-mail addresses: andreas.pfeiffer@dlr.de (A. Pfeiffer), arnold@mathematik.uni-halle.de (M. Arnold).

Long-term simulations of nonlinear problems and the time discretisation of partial differential equations have, however, shown that the highly developed classical time integration methods are often less robust and less efficient than rather simple nonstandard methods that consider explicitly structural properties of the analytical solution like symmetry and invariants [8].

The present paper contributes to a deeper understanding of such a new class of methods that was presented recently in the engineering literature. We show that the new approach may be combined successfully with classical high-order collocation and Nyström methods. As one result we get symmetric and reversible integration methods of arbitrary high order which conserve energy and linear and angular momentum in the time integration of Hamiltonian systems.

The paper goes back to the work of Betsch and Steinmann [5] who modify the distance potential in an N -body problem such that the system's total energy is preserved under time discretisation by a classical Galerkin type method.

After this short introduction the paper is organised as follows. In Section 2 we recall some basic properties of Hamiltonian systems and illustrate by an example from celestial mechanics that classical time integration methods often fail to preserve structural properties of Hamiltonian systems like energy conservation or conservation of linear and angular momentum.

In Section 3 perturbed potential functions are used as general construction principle for energy conserving Galerkin and collocation methods that may both be considered as special cases of the new class of generalised Nyström methods. Energy conservation in the numerical solution defines implicitly the perturbation of the potential.

Under mild assumptions we show in Section 4 that the perturbation parameter κ and the numerical solution are uniquely determined by the condition for energy conservation. Furthermore, the perturbation does not affect the order of convergence. That means that energy conservation is achieved without loss of accuracy.

Other structural properties of the numerical solution are studied in Section 5. Conditions for the preservation of linear and quadratic invariants are given. The symmetry and the reversibility of a modified Nyström method are equivalent to a set of conditions that is always satisfied if the corresponding classical Nyström method is symmetric and reversible.

The paper is completed by the presentation of numerical test results in Section 6 and a summary in Section 7. The numerical tests illustrate the order results of Section 4. Perturbed potentials are compared with classical projection techniques showing clearly that the novel approach is superior, especially in long-term simulations.

2. Hamiltonian systems

Many technical or physical models are represented by Hamiltonian systems. The motion of a mechanical N -body problem may be considered as a typical example of a Hamiltonian system that is described by a classical system of ordinary differential equations. Furthermore, many Hamiltonian systems result from space discretisation of time dependent partial differential equations in the field of elastodynamics [6].

In both cases the solution trajectories have some geometrical properties such as energy conservation and often also conservation of linear and angular momentum. The flow produced by the solution curves is symmetric and reversible. Reversibility means, that a reversed initial velocity does not change the solution trajectory, but only its direction.

In classical mechanics the Hamiltonian principle is used as a variational principle to describe the dynamics of the mechanical system. The corresponding system of ordinary differential equations has a special structure resulting from the Hamiltonian function. Often the Hamiltonian function is the sum of a potential $V(x)$ and the kinetic energy $K(x, p) = \frac{1}{2}p^\top M(x)^{-1}p$ with a mass matrix $M(x)$ [2, Chapter 3.4]. Here, x is the generalised coordinate, p stands for the generalised momentum.

Definition 2.1 (*Hamiltonian system*). We assume that $V: \Omega \rightarrow \mathbb{R}$ and $M: \Omega \rightarrow \mathbb{R}^{d \times d}$ are smooth maps on an open set $\Omega \subset \mathbb{R}^d$. Furthermore, $M(x)$ is symmetric positive definite for $x \in \Omega$. Then we call the function $H: S \rightarrow \mathbb{R}$,

$$H(x, p) := V(x) + K(x, p) := V(x) + \frac{1}{2}p^\top M(x)^{-1}p$$

Hamiltonian function on $S := \Omega \times \mathbb{R}^d$ with the associated Hamiltonian system

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} H_p^\top(x, p) \\ -H_x^\top(x, p) \end{pmatrix}.$$

With $y := M(x)^{-1}p$ the Hamiltonian system gets the following Lagrangian form:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} y \\ g(x, y) \end{pmatrix}, \quad (1)$$

with $g(x, y) := M(x)^{-1}(-\nabla V(x) + g_2(x, y) - g_1(x, y))$,

$$g_1(x, y) := \begin{pmatrix} y^\top M'_1(x)y \\ \vdots \\ y^\top M'_d(x)y \end{pmatrix} \quad \text{and} \quad g_2(x, y) := \frac{1}{2} \begin{pmatrix} y^\top M_{x_1}(x)y \\ \vdots \\ y^\top M_{x_d}(x)y \end{pmatrix}.$$

M'_i represents the matrix of derivatives of the i th column of M , and M_{x_i} is the partial derivative of M w.r.t. x_i . We define the *energy function* E as the Hamiltonian H in the new coordinates x, y :

$$E(x, y) := H(x, M(x)y) = V(x) + \frac{1}{2}y^\top M(x)y.$$

It is well known, that the Hamiltonian H is a first integral of the associated Hamiltonian system. Therefore the energy E is constant along the solution trajectories. The special case, that the kinetic energy K depends on p only, motivates the following definition.

Definition 2.2 (*Separable Hamiltonian system*). The Hamiltonian function $H(x, p) = V(x) + K(p) = V(x) + \frac{1}{2}p^\top M^{-1}p$ with constant mass matrix M is called *separable Hamiltonian function* with the *separable Hamiltonian system*

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} M^{-1}p \\ -\nabla V(x) \end{pmatrix}. \quad (2)$$

The N -body problem is a typical example of a separable Hamiltonian system that is characterised by potential forces which depend only on the distances between the bodies [5,8].

Definition 2.3 (*Distance potential*). We partition $x = (x^{1^\top}, \dots, x^{N^\top})^\top$ by $x^i \in \mathbb{R}^m$ ($m := d/N \in \mathbb{N}$) and assume that V_{ij} ($i > j$) are smooth functions on open sets in \mathbb{R} . Then the potential $V : \Omega \rightarrow \mathbb{R}$,

$$V(x) := \sum_{i=2}^N \sum_{j=1}^{i-1} V_{ij}(\|x^i - x^j\|_2),$$

is called *distance potential*.

Theorem 2.4 (First integrals for distance potentials). *For distance potentials the following functions are first integrals of the separable Hamiltonian system in its Lagrangian form (1):*

- (a) $F(x, y) := c^\top M y$ with $c^\top := (\bar{c}^\top, \dots, \bar{c}^\top)$, $\bar{c} \in \mathbb{R}^m$,
- (b) $F(x, y) := x^\top C M y$ with $C := \text{blockdiag}(\bar{C}, \dots, \bar{C})$, $\bar{C} = -\bar{C}^\top \in \mathbb{R}^{m \times m}$, if CM is skew symmetric: $(CM)^\top = -CM$.

Proof. For separable Hamiltonian systems we have $g(x, y) = -M^{-1} \nabla V(x)$ and obtain $F_x(x, y)y - F_y(x, y)M^{-1} \nabla V(x) = 0$ for $(x, y) \in S$. \square

As a consequence separable Hamiltonian systems (2) with distance potentials ($m = 3$) and a diagonal mass matrix M have the following first integrals: linear momentum $\sum_{i=1}^N p^i$ and angular momentum $\sum_{i=1}^N x^i \times p^i$ [8, Example IV.1.3].

Important structural properties of Hamiltonian systems can be characterised by the *flow*, which is symmetric and reversible. It is well known, that the preservation of these structural properties and the preservation of invariants in Hamiltonian systems yield good results in long-term simulations, see [8, Chapters V, XI].

In the following we consider one-step methods $(x_1, y_1) = \Phi(h, (x_0, y_0))$ that define the *numerical flow* $\Phi : I \times Q \rightarrow S$ on an open interval $I \subset \mathbb{R}$ with $0 \in I$ and an open set $Q \subset S$. It turns out that numerical integration methods do in general not fulfil the following identities:

symmetry: $\Phi(-h, \Phi(h, (x, y))) = (x, y)$ and

reversibility: $\Phi(h, \rho(\Phi(h, (x, y)))) = \rho(x, y)$ with $\rho(x, y) := (x, -y)$.

Furthermore, invariants of Hamiltonian systems are in general not preserved by numerical methods. For the purpose of illustration we integrate a model for the motion of the outer planets in our Solar System [8, Section I.2.3]. As integration method we choose the Radau IIA method with $s = 3$ stages and a constant step size of $h = 100$ days. The corresponding Hamiltonian system is separable and has the invariants energy, linear and angular momentum. Fig. 1 shows that only the linear momentum is conserved up to machine precision. Angular momentum and energy have considerable deflections.

A classical way to enforce the conservation of an invariant is to project in each integration step the numerical solution $(\tilde{x}_1, \tilde{y}_1)$ onto the manifold being defined by the invariant F . The solution (x_1, y_1) is chosen as a solution of the following minimisation problem:

$$\min_{(x_1, y_1) \in S} \|(x_1, y_1) - (\tilde{x}_1, \tilde{y}_1)\|_2, \quad \text{with } F(x_1, y_1) = F(x_0, y_0).$$

(x_0, y_0) denotes the solution of the previous integration step. Hairer et al. [8, Chapter IV.4] show that the projection does not decrease the order of convergence, but in general a symmetric method is changed

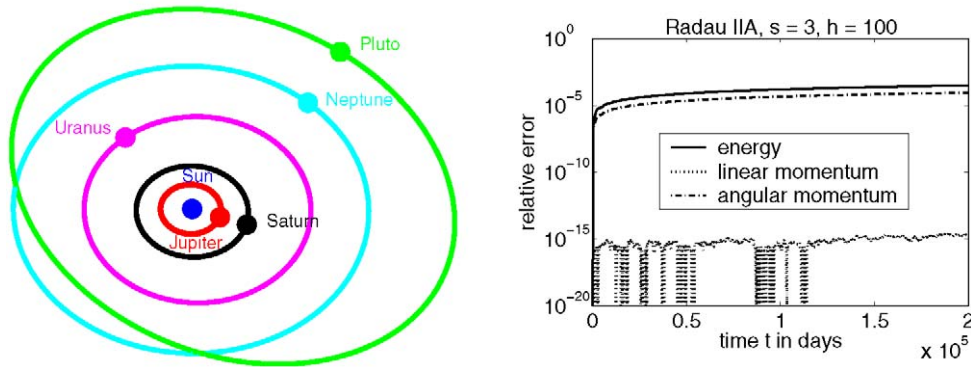


Fig. 1. Simulation of the outer solar system over 200 000 days with a Radau IIA method.

to an unsymmetric one [8, Section V.4.1]. On the other hand, some integration methods preserve certain invariants, which are not conserved after a projection on an additional invariant [1]. Therefore, in the next section we try to keep favorable properties of well-known methods, if they are modified to guarantee energy conservation.

3. Modified Nyström methods

Numerical methods for ordinary differential equations provide in general only approximations to the solution trajectories. It is therefore natural to look for numerical solutions that share all essential structural properties of the analytical solution. In the present section a class of one-step methods for Hamiltonian systems is introduced that guarantees conservation of the total energy.

The integration interval $[0, T]$ is discretised by a time grid $\{t_0, \dots, t_N\}$ with $0 = t_0 < t_1 < \dots < t_N = T$ and time step sizes $h_n := t_{n+1} - t_n$ for $n = 0, 1, \dots, N - 1$. It is sufficient to consider only the step from $t_0 := 0$ to $t_1 = h_1 =: h$ and to use as initial value (x_0, y_0) the numerical solution from the previous step. We approximate the solution on $[0, h]$ by polynomials (\bar{x}, \bar{y}) up to degree $s \in \mathbb{N}$ and set $(x_1, y_1) := (\bar{x}(h), \bar{y}(h))$. At $t = 0$ the polynomials \bar{x}, \bar{y} shall satisfy $(\bar{x}(0), \bar{y}(0)) = (x_0, y_0)$.

Finite element methods for solving ordinary differential equations are frequently considered in the engineering literature [5], because the ordinary differential equations have often its origin in partial differential equations that are semi-discretised in space by finite elements [6]. We follow the idea of finite elements for ordinary differential equations and rewrite the resulting methods in a generalised Runge–Kutta form.

The starting point of the Galerkin approach is the weak formulation of the Hamiltonian system with test functions v :

$$\int_0^h (\dot{\bar{x}}(t) - \bar{y}(t)) v(t) dt = 0, \quad (3a)$$

$$\int_0^h (\dot{\bar{y}}(t) - g(\bar{x}(t), \bar{y}(t))) v(t) dt = 0. \quad (3b)$$

We choose test functions $v \in \mathbb{P}_{s-1}$, this means polynomials up to degree $s - 1$. The integrals in (3) may be solved exactly except

$$\int_0^h g(\bar{x}(t), \bar{y}(t)) v(t) dt. \quad (4)$$

Betsch and Steinmann [5] suggest to use quadrature formulas to solve this integral. For separable Hamiltonian systems they present modifications of the quadrature formulas in order to preserve energy, linear and angular momentum. The results are illustrated by numerical examples with low order methods.

In the following we generalise one of these modified methods of Betsch and Steinmann and show in Sections 4 and 5, that this approach yields a class of energy conserving methods including methods of arbitrary high order and methods that preserve additionally a number of other structural properties.

To approximate (4) we use a quadrature formula with nodes $c_1, \dots, c_s \neq 0$ and weights $w_1, \dots, w_s \neq 0$ that has the minimum order s . Additionally, we insert a scalar $\kappa \in \mathbb{R}$ as a perturbation of the vector field g :

$$\int_0^h g(\bar{x}(t), \bar{y}(t)) v(t) dt \approx h \sum_{k=1}^s w_k \kappa g(\bar{x}(hc_k), \bar{y}(hc_k)) v(hc_k).$$

For test functions $v \in \mathbb{P}_{s-1}$ a quadrature formula with s nodes is chosen to allow a transformation of the methods into Nyström form, which is for $\kappa := 1$ well known in the field of numerical solution methods for second order differential equations.

Modified Galerkin approach

For a quadrature formula with nodes $c_1, \dots, c_s \in (0, 1]$ and corresponding weights $w_1, \dots, w_s \neq 0$ the polynomials $(\bar{x}, \bar{y}) \in \mathbb{P}_s^{2d}$ with $(\bar{x}(0), \bar{y}(0)) = (x_0, y_0)$ have to satisfy

$$\int_0^h (\dot{\bar{x}}(t) - \bar{y}(t)) v(t) dt = 0, \quad (5a)$$

$$\int_0^h \dot{\bar{y}}(t) v(t) dt - h \sum_{k=1}^s w_k \kappa g(\bar{x}(hc_k), \bar{y}(hc_k)) v(hc_k) = 0, \quad (5b)$$

for all $v \in \mathbb{P}_{s-1}$ and a fixed $\kappa \in \mathbb{R}$.

Existence and uniqueness of these polynomials \bar{x} and \bar{y} will be considered in Section 4. Conditions will be given to guarantee $\bar{x}(hc_k) \in \Omega$ for $k = 1, \dots, s$.

In the field of Runge–Kutta methods collocation approaches play a fundamental role in constructing new methods [9, Chapter II.7]. Now we will show, that a modified collocation idea yields energy conserving methods for Hamiltonian systems, which are very closely related to the modified Galerkin method (5).

Modified collocation approach

Consider a quadrature formula with minimum order s , nodes $c_1, \dots, c_s \in [0, 1]$ and weights $w_1, \dots, w_s \neq 0$. We consider polynomials (\bar{x}, \bar{y}) that satisfy the modified collocation conditions

$$\dot{\bar{x}}(hc_i) = \bar{y}(hc_i), \quad (6a)$$

$$\dot{\bar{y}}(hc_i) = \kappa g(\bar{x}(hc_i), \bar{y}(hc_i)), \quad (6b)$$

for $i = 1, \dots, s$ and a given $\kappa \in \mathbb{R}$.

Betsch and Steinmann [4] point out that the Galerkin approach (5) and the collocation approach (6) are equivalent for $s = 2, 3$ and Gauss quadrature. This result may be generalised to all modified methods with Gauss nodes c_i :

Theorem 3.1 (Equivalence for Gauss methods). *For a fixed $\kappa \in \mathbb{R}$ and the Gauss nodes c_1, \dots, c_s with the corresponding weights w_1, \dots, w_s the modified Galerkin approach (5) and the modified collocation approach (6) are equivalent.*

Proof. The proof is based on the high order $2s$ of Gauss quadrature [7, Chapter 9.3]. Therefore all polynomials $u \in \mathbb{P}_{2s-1}^d$ fulfil the equation

$$\int_0^h u(t) dt = h \sum_{k=1}^s w_k u(hc_k). \quad (*)$$

(a) We assume, that the collocation conditions (6) hold. Then all test polynomials $v \in \mathbb{P}_{s-1}$ satisfy:

$$\begin{aligned} \int_0^h (\dot{\bar{x}}(t) - \bar{y}(t)) v(t) dt &\stackrel{(*)}{=} h \sum_{k=1}^s w_k \underbrace{(\dot{\bar{x}}(hc_k) - \bar{y}(hc_k))}_{=0} v(hc_k) = 0, \\ \int_0^h \dot{\bar{y}}(t) v(t) dt &\stackrel{(*)}{=} h \sum_{k=1}^s w_k \dot{\bar{y}}(hc_k) v(hc_k) \stackrel{(6b)}{=} h\kappa \sum_{k=1}^s w_k g(\bar{x}(hc_k), \bar{y}(hc_k)) v(hc_k). \end{aligned}$$

Thus (\bar{x}, \bar{y}) fulfil the Galerkin conditions (5).

(b) Suppose that the polynomials (\bar{x}, \bar{y}) satisfy (5). We choose test polynomials $v := \bar{l}_i \in \mathbb{P}_{s-1}$ with $\bar{l}_i(h\tau) := l_i(\tau) := \prod_{j=1, j \neq i}^s \frac{\tau - c_j}{c_i - c_j}$ (for $\tau \in \mathbb{R}$) and obtain:

$$\begin{aligned} 0 &\stackrel{(5a)}{=} \int_0^h (\dot{\bar{x}}(t) - \bar{y}(t)) \bar{l}_i(t) dt \stackrel{(*)}{=} h \sum_{k=1}^s w_k (\dot{\bar{x}}(hc_k) - \bar{y}(hc_k)) \delta_{ik} = h w_i (\dot{\bar{x}}(hc_i) - \bar{y}(hc_i)), \\ h\kappa w_i g(\bar{x}(hc_i), \bar{y}(hc_i)) &= h\kappa \sum_{k=1}^s w_k g(\bar{x}(hc_k), \bar{y}(hc_k)) \bar{l}_i(hc_k) \\ &\stackrel{(5b)}{=} \int_0^h \dot{\bar{y}}(t) \bar{l}_i(t) dt \stackrel{(*)}{=} h \sum_{k=1}^s w_k \dot{\bar{y}}(hc_k) \delta_{ik} = h w_i \dot{\bar{y}}(hc_i). \end{aligned}$$

Because of $hw_i \neq 0$ Eq. (6) is obtained. \square

Modified Galerkin methods (5) and modified collocation methods (6) are applied to a system with very special right-hand side that allows to eliminate the node values $\bar{x}(hc_i)$ analytically. The resulting Runge–Kutta like representation of (5) and (6) may be considered as a generalisation of the well-known Nyström methods for second order ordinary differential equations.

Theorem 3.2 (Nyström representation). *We define*

$$X := \begin{pmatrix} X_1 \\ \vdots \\ X_s \end{pmatrix}, \quad Y := \begin{pmatrix} Y_1 \\ \vdots \\ Y_s \end{pmatrix}, \quad X^0 := \mathbb{1} \otimes x_0, \quad Y^0 := \mathbb{1} \otimes y_0,$$

with $\mathbb{1} := (1, \dots, 1)^T \in \mathbb{R}^s$ and

$$G(X, Y) := \begin{pmatrix} g(X_1, Y_1) \\ \vdots \\ g(X_s, Y_s) \end{pmatrix}, \quad \Upsilon := \begin{pmatrix} \Upsilon_1 \\ \vdots \\ \Upsilon_s \end{pmatrix} \in \mathbb{R}^{ds}, \quad X_i \in \Omega, \quad Y_i, \quad \Upsilon_i \in \mathbb{R}^d.$$

Then, both the modified Galerkin approach (5) and the modified collocation approach (6) can be transformed into the following Nyström form:

$$\Upsilon = G(X^0 + hCY^0 + h^2\kappa A\Upsilon, Y^0 + h\kappa \bar{A}\Upsilon), \quad (7a)$$

$$x_1 = x_0 + hy_0 + h^2\kappa \sum_{i=1}^s b_i \Upsilon_i, \quad (7b)$$

$$y_1 = y_0 + h\kappa \sum_{i=1}^s w_i \Upsilon_i. \quad (7c)$$

This representation contains the coefficient matrices $C := c \otimes I_d$, $A := a \otimes I_d$, $\bar{A} := \bar{a} \otimes I_d$ and $c := \text{diag}(c_1, \dots, c_s)$, $a, \bar{a} \in \mathbb{R}^{s \times s}$. The matrices a, \bar{a} and the coefficients $b_i \in \mathbb{R}$ are determined by the nodes c_1, \dots, c_s and the weights w_1, \dots, w_s :

(a) Galerkin approach.

$$a := \hat{a}\bar{a}, \quad \bar{a} := \gamma^{-1}w, \quad b_i := \sum_{j=1}^s L_j(1)a_{ji},$$

with $w := \text{diag}(w_1, \dots, w_s)$, $\hat{a} := \gamma^{-1}\theta$ and

$$\gamma := (\gamma_{ij})_{i,j=1}^s = \left(\int_0^1 \dot{L}_j(\tau) l_i(\tau) d\tau \right), \quad \theta := (\theta_{ij})_{i,j=1}^s = \left(\int_0^1 L_j(\tau) l_i(\tau) d\tau \right),$$

$$l_i(\tau) := \prod_{\substack{j=1 \\ j \neq i}}^s \frac{\tau - c_j}{c_i - c_j}, \quad L_i(\tau) := \prod_{\substack{j=0 \\ j \neq i}}^s \frac{\tau - c_j}{c_i - c_j}, \quad c_0 := 0.$$

(b) *Collocation approach.*

$$a := \bar{a}^2, \quad b^T := (w_1, \dots, w_s) \bar{a}, \quad \bar{a} := (\bar{a}_{ij})_{i,j=1}^s := \left(\int_0^{c_i} l_j(\tau) d\tau \right),$$

$$\text{with } l_i(\tau) := \prod_{j=1, j \neq i}^s \frac{\tau - c_j}{c_i - c_j}.$$

Proof. (a) In the Galerkin approach we use a representation of (\bar{x}, \bar{y}) by means of the Lagrangian basis polynomials L_i and use in (5) test polynomials $v := l_i$. The resulting system of linear equations may be solved w.r.t. $(\bar{x}(hc_i), \bar{y}(hc_i))$ by some straightforward computations, see also [10, Theorem 3.14].

(b) follows directly from classical results for collocation, Runge–Kutta and Nyström methods, see, e.g., [8, Theorem II.1.4]. \square

The Nyström representation (7) of the constructed methods motivates to modify standard Nyström methods by introducing the scalar κ . In the following we will use this generalised Nyström form without paying attention to any specific way for obtaining the coefficients a, \bar{a}, b .

In the engineering literature methods that conserve energy [5,6] and methods with a prescribed energy decay [3] found special interest. Energy conservation or energy decay define implicitly the scalar κ . In the following we will consider in more detail energy conserving methods. They are characterized by:

$$E(x_1, y_1) = E(x_0, y_0). \quad (8)$$

Inserting the right-hand sides of Eqs. (7b) and (7c) into (8) leads to an implicit and coupled system of nonlinear equations (7a), (8) for the unknowns \mathcal{Y} and κ , which has dimension $ds + 1$.

4. Properties of the numerical solution

After defining the modified Nyström methods in Section 3 we investigate now existence and uniqueness of the numerical solution and its structural properties.

Definition 4.1 (*Regular initial value*). The initial value $(x_0, y_0) \in S$ for a Hamiltonian system is called *regular*, if

$$y_0^T (\nabla V(x_0) + g_2(x_0, y_0)) \neq 0,$$

and *singular* otherwise.

Theorem 4.2 (Existence and uniqueness of the numerical solution). *We consider a consistent generalised Nyström method (7) for Hamiltonian systems that satisfies the energy equation (8).*

For each regular initial value $(x_0, y_0) \in S$ there is a neighbourhood $Q \subset S$ of (x_0, y_0) and a constant $\bar{h} > 0$ such that Eqs. (7), (8) with $|h| \leq \bar{h}$ and initial values $(\hat{x}, \hat{y}) \in Q$ have a unique solution in Q that depends continuously differentiable on h and (\hat{x}, \hat{y}) .

Proof. We write the system of Eqs. (7), (8) as $R(\mathcal{Y}, h, \kappa, \hat{x}, \hat{y}) = 0$ with the map

$$R(\mathcal{Y}, h, \kappa, \hat{x}, \hat{y}) := \begin{pmatrix} \mathcal{Y} - G(\hat{X}^0 + hC\hat{Y}^0 + h^2\kappa A\mathcal{Y}, \hat{Y}^0 + h\bar{A}\mathcal{Y}) \\ r(\mathcal{Y}, h, \kappa, \hat{x}, \hat{y}) \end{pmatrix}, \quad \hat{X}^0 := \mathbb{1} \otimes \hat{x}, \quad \hat{Y}^0 := \mathbb{1} \otimes \hat{y},$$

and

$$\begin{aligned} r(\Upsilon, h, \kappa, \hat{x}, \hat{y}) &:= \frac{1}{h} \left[E \left(\hat{x} + h\hat{y} + h^2\kappa \sum_{i=1}^s b_i \Upsilon_i, \hat{y} + h\kappa \sum_{i=1}^s w_i \Upsilon_i \right) - E(\hat{x}, \hat{y}) \right] \\ &= \frac{1}{h} [E(z(1)) - E(z(0))], \end{aligned}$$

with $z(\theta) := (\hat{x} + \theta h\hat{y} + \theta h^2\kappa \sum_{i=1}^s b_i \Upsilon_i, \hat{y} + \theta h\kappa \sum_{i=1}^s w_i \Upsilon_i)$ ($\theta \in [0, 1]$). We get

$$r = \frac{1}{h} \int_0^1 \frac{dE(z(\theta))}{d\theta} d\theta = \int_0^1 \left[E_x(z(\theta)) \left(\hat{y} + h\kappa \sum_{i=1}^s b_i \Upsilon_i \right) + E_y(z(\theta)) \kappa \sum_{i=1}^s w_i \Upsilon_i \right] d\theta.$$

Therefore the function r is continuously differentiable in a neighbourhood of $(G(\hat{X}^0, \hat{Y}^0), 0, 1, x_0, y_0)$ and has the partial derivative

$$r_\kappa(G(\hat{X}^0, \hat{Y}^0), 0, 1, x_0, y_0) = -y_0^\top (\nabla V(x_0) + g_2(x_0, y_0)) =: \alpha.$$

For regular initial values (x_0, y_0) we have $\alpha \neq 0$ and the Jacobian matrix

$$R_{\Upsilon, \kappa}(G(\hat{X}^0, \hat{Y}^0), 0, 1, x_0, y_0) = \begin{pmatrix} I_{ds} & 0 \\ * & \alpha \end{pmatrix}$$

is nonsingular. The Implicit function theorem guarantees, that there exist smooth functions $\Upsilon(h, \hat{x}, \hat{y})$, $\kappa(h, \hat{x}, \hat{y})$ with $R(\Upsilon(h, \hat{x}, \hat{y}), h, \kappa(h, \hat{x}, \hat{y}), \hat{x}, \hat{y}) = 0$, if $|h|$ is sufficiently small. \square

In [10] singular initial values with $y_0 = 0$, $\nabla V(x_0) \neq 0$ are investigated in more detail and the existence of a continuously differentiable solution $(x(h), y(h))$ of (7) and (8) is shown for sufficiently small $|h|$.

Standard collocation methods have the same order of convergence as the underlying quadrature formula [8, Section II.1.2]. In the following we consider the question if a similar *superconvergence* result can be achieved by the Galerkin methods (5) at least with $\kappa := 1$. Furthermore we study the influence of the energy equation (8) on the order of convergence of the modified Galerkin methods and of the modified general Nyström methods.

First, we investigate the Galerkin approach (5) with quadrature formulas of high order and with fixed $\kappa := 1$. Similar to the proof of the convergence theorem [9, Theorem II.7.9] for collocation methods we show that the local error has at least order $s + 1$ ($s \dots$ number of quadrature nodes). Further, the derivatives of the polynomials (\bar{x}, \bar{y}) approximate the derivatives of the analytical solution. The order $s + 1$ is decreased by the order of the derivative (Lemma 4.4). Both results are valid for collocation methods, too. Finally, this is sufficient to prove the superconvergence for the Galerkin methods (Theorem 4.5).

Lemma 4.3 (Order s for the Galerkin approach). *We consider a problem*

$$\dot{z} := \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} y \\ g(x, y) \end{pmatrix} =: f(z),$$

with a smooth vector field g and a quadrature formula with s nodes, which has the order $p \geq 2s - 1$. For an initial value $z_0 = (x_0, y_0) \in S$ and sufficiently small $|h|$ the numerical approximation of the Galerkin approach with $\kappa := 1$ is denoted by $\bar{z} = (\bar{x}, \bar{y})$ and the analytical solution by $z = (x, y)$ with $\bar{z}(0) = z(0) = z_0$.

Then the local error is bounded by:

$$\max_{t \in [0, h]} \|\bar{z}(t) - z(t)\| = \mathcal{O}(h^{s+1}).$$

Proof. We write (7) as $R(x_1, y_1) = 0$ with a suitable map R and insert the analytical solution $z = (x, y)$ into this system. Then we conclude: $R(z(h)) = \mathcal{O}(h^{p-s+2}) + \mathcal{O}(h^{s+1}) = \mathcal{O}(h^{s+1})$, because $p \geq 2s - 1$ is assumed. This estimate implies the proposition since the derivatives of the inverse map R^{-1} are bounded, see [10, Lemma 3.26] for more details. \square

Lemma 4.4 (Convergence of the derivatives). *Under the assumptions of Lemma 4.3 the following estimates hold for $k = 0, \dots, s$:*

$$\max_{t \in [0, h]} \|\bar{z}^{(k)}(t) - z^{(k)}(t)\| = \mathcal{O}(h^{s+1-k}).$$

Proof. The lemma can be proven like the same proposition for collocation methods [9, Theorem II.7.10]. \square

Theorem 4.5 (Superconvergence for Galerkin methods). *The assumptions of Lemma 4.3 shall be fulfilled. Then the Galerkin method has the same order p as the underlying quadrature formula.*

Proof. With the Galerkin polynomial \bar{z} we define the defect $\delta(t, \bar{z}) := \dot{\bar{z}}(t) - f(\bar{z})$ for $t \in [0, h]$ and $\bar{z} \in S$. Then \bar{z} is the solution of the following perturbed initial value problem for $t \in [0, h]$:

$$\dot{\bar{z}}(t) = f(\bar{z}) + \delta(t, \bar{z}), \quad \bar{z}(0) = z(0) = z_0.$$

The theorem of Gröbner and Alekseev [9, Theorem I.14.5] states

$$\bar{z}(h) - z(h) = \int_0^h \Psi(t) dt \tag{*}$$

with $\Psi(t) := D(t)\delta(t, \bar{z}(t))$ and the map $D(t) := \partial z / \partial \bar{z}(h, t, \bar{z}(t))$. Here the analytical solution z is considered as a function $z(h, t, \bar{z})$ depending on the final time point h , initial time point t and the initial value \bar{z} at time t (!). Because the derivatives $\bar{z}^{(k)}$ are bounded (Lemma 4.4), we can conclude with the Taylor polynomial $T_s D(t) := \sum_{k=0}^s D^{(k)}(0)t^k/k!$:

$$D(t) = T_{s-1} D(t) + \mathcal{O}(h^s) \quad \text{for } t \in [0, h].$$

Applying Lemma 4.4 we get the estimate $\delta(t, \bar{z}(t)) = \mathcal{O}(h^s)$ and

$$\Psi(t) = T_{s-1} D(t)\delta(t, \bar{z}(t)) + \mathcal{O}(h^{2s}).$$

Now we can write (*) as

$$\bar{z}(h) - z(h) = \int_0^h T_{s-1} D(t)\delta(t, \bar{z}(t)) dt + \mathcal{O}(h^{2s+1}).$$

To estimate this integral, we consider an arbitrary polynomial $v \in \mathbb{P}_{s-1}$. Because of the Galerkin conditions (5), the order p of the quadrature formula and the boundedness of the derivatives of \bar{z} we get the following equations:

$$\int_0^h (\dot{\bar{x}}(t) - \bar{y}(t))v(t) dt = 0,$$

$$\int_0^h (\dot{\bar{y}}(t) - g(\bar{z}(t)))v(t) dt = \int_0^h \dot{\bar{y}}(t)v(t) dt - h \sum_{k=1}^s w_k g(\bar{z}(hc_k))v(hc_k) + \mathcal{O}(h^{p+1}) = \mathcal{O}(h^{p+1}).$$

We obtain

$$\int_0^h v_h(t)\delta(t, \bar{z}(t)) dt = \mathcal{O}(h^{p+1}),$$

for polynomials $v_h \in \mathbb{P}_{s-1}$, which depend on h , but have bounded derivatives for $h \rightarrow 0$. Each row of $T_{s-1}D$ consists of such polynomials (due to the definition of D and the boundedness of the derivatives of \bar{z}), therefore we conclude

$$\int_0^h T_{s-1}D(t)\delta(t, \bar{z}(t)) dt = \mathcal{O}(h^{p+1}),$$

and finally

$$\bar{z}(h) - z(h) = \mathcal{O}(h^{p+1}) + \mathcal{O}(h^{2s+1}) = \mathcal{O}(h^{p+1}), \quad \text{since } p \leq 2s. \quad \square$$

There remains the open question, which order of convergence we can achieve, if the order of the quadrature formula is less than $2s - 1$. If one can prove the statement of Lemma 4.3 for quadrature formulas with order less than $2s - 1$, then Theorem 4.5 is proved as well for the corresponding Galerkin methods. The assumption $p \geq 2s - 1$ is only used for the proof of Lemma 4.3.

For classical projection methods [8, Chapter IV.4] the order of convergence is identical to the order of the underlying method without projection. In Theorem 4.6 we obtain the same result for the modified Nyström methods (7) with the energy equation (8).

Theorem 4.6 (Convergence with energy equation). *We suppose that a generalised Nyström method (7) with $\kappa := 1$ has order p of convergence. Then the modified Nyström method with κ being determined implicitly by the energy equation (8) has the same order p , if the solution remains inside a set without singular initial values.*

Proof. For a regular initial value $z_0 := (x_0, y_0)$ with $z(0) = z_1(0, \kappa) = z_0$ the analytical solution shall be denoted by z and the unique, continuous solution of the Nyström equations (7) in the neighbourhood of $(0, 1) \subset \mathbb{R}^2$ by $z_1(h, \kappa)$, see Theorem 4.2. Due to the assumption on the general Nyström method it holds

$$z_1(h, 1) = z(h) + \mathcal{O}(h^{p+1}). \quad (*)$$

As in the proof of Theorem 4.2 we formulate the energy condition as

$$r(h, \kappa) := \frac{1}{h} [E(z_1(h, \kappa)) - E(z_0)] = 0.$$

Since $E(z(h)) = E(z_0)$ and (*) we estimate

$$r(h, 1) = h^{-1} \mathcal{O}(h^{p+1}) = \mathcal{O}(h^p).$$

Because of $r_\kappa(0, 1) = -y_0^\top (\nabla V(x_0) + g_2(x_0, y_0)) \neq 0$ we have $|r_\kappa(h, \kappa)| > 0$ in a neighbourhood of $(0, 1)$, therefore $|\kappa(h) - 1| = \mathcal{O}(h^p)$. From (7b) and (7c) we obtain the final estimate:

$$\|\bar{z}(h) - z(h)\| \leq \|z_1(h, \kappa(h)) - z_1(h, 1)\| + \|z_1(h, 1) - z(h)\| = \mathcal{O}(h^{p+1}). \quad \square$$

5. Preservation of structural properties

The above constructed modified Nyström methods are energy conserving for Hamiltonian systems. Additional invariants like linear and angular momentum can be preserved under conditions on the coefficients (Theorem 5.1). We want to transfer symmetry and reversibility of the analytical flow to the numerical solution because of the advantages in long-term simulations, see Section 2. Necessary and sufficient criteria for the symmetry of Runge–Kutta methods can be found in [8, Chapter V.2] and [9, Chapter II.8]. We extend these investigations to Nyström methods and get in Theorem 5.3 similar conditions on the coefficients. With these conditions we are able to prove symmetry and reversibility of the energy conserving methods in Theorem 5.4.

Theorem 5.1 (Preservation of linear and quadratic invariants).

- (a) *The modified Nyström method (7) preserves linear invariants depending only on y .*
 (b) *If the coefficients fulfil for $i, j = 1, \dots, s$ the conditions*

$$\begin{aligned} b_i &= w_i(1 - c_i), \\ w_i(b_j - a_{ij}) &= w_j(b_i - a_{ji}), \end{aligned}$$

then quadratic first integrals of the form $F(x, y) = x^\top C y$ with a skew symmetric matrix $C = -C^\top \in \mathbb{R}^{d \times d}$ are preserved.

Proof. (a) Let $F(x, y) := c^\top y$ denote the linear invariant with $c \in \mathbb{R}^d$. Then for all $(x, y) \in S$ the equation

$$c^\top g(x, y) = 0$$

holds and we get from (7c):

$$c^\top y_1 = c^\top y_0 + h\kappa \sum_{i=1}^s w_i \underbrace{c^\top \gamma_i}_{=0} = c^\top y_0.$$

- (b) The proof for the quadratic invariants follows the proof of Theorem IV.2.5 in [8]. \square

Corollary 5.2 (Preservation of linear and angular momentum). *Under the assumptions of Theorem 5.1 the modified Nyström method preserves the linear and angular momentum.*

The definition of symmetry in Section 2 uses properties of the numerical flow. To investigate the symmetry of the modified Nyström method we look for conditions on the coefficients and extend the corresponding Theorem II.8.8 from [9], see also [10, Theorem 3.47].

Theorem 5.3 (Conditions for symmetry). *We consider a consistent generalised Nyström method (7) with $\kappa := 1$, $w_i \neq 0$ for $i = 1, \dots, s$ and pairwise different nodes $c_i := \sum_{j=1}^s \bar{a}_{ij}$ that satisfy $c_1 < c_2 < \dots < c_s$.*

Then the method is symmetric, if and only if the coefficients fulfil the following conditions for $i, j = 1, \dots, s$:

$$c_i = 1 - c_{s+1-i},$$

$$w_i = w_{s+1-i},$$

$$b_i = w_i - b_{s+1-i},$$

$$\bar{a}_{ij} = w_j - \bar{a}_{s+1-i, s+1-j},$$

$$a_{ij} = b_j - c_{s+1-i} w_j + a_{s+1-i, s+1-j}.$$

Theorem 5.4 (Symmetry and reversibility). *If the coefficients fulfil the conditions of Theorem 5.3, then the modified Nyström method (7) with the energy equation (8) is symmetric and reversible for Hamiltonian systems.*

Proof. To prove the symmetry we use Theorem 5.3 and the fact, that the energy equation $E(x_1, y_1) = E(x_0, y_0) \iff E(x_0, y_0) = E(x_1, y_1)$ is symmetric.

A simple calculation for the modified Nyström methods with energy equation shows the identity

$$\rho(\Phi(h, (x_0, y_0))) = (x_1, -y_1) = \Phi(-h, \rho(x_0, y_0)).$$

Theorem V.1.5 in [8] proves then the equivalence between symmetry and reversibility. \square

The Galerkin and collocation methods are symmetric and reversible, if the nodes of the quadrature formula are symmetric in terms of $c_i = 1 - c_{s+1-i}$ for $i = 1, \dots, s$ [10, Theorems 3.46, 3.49].

This result completes the theoretical analysis of generalised Nyström methods that started in Section 3 with the construction of energy conserving Nyström methods for the numerical solution of Hamiltonian systems in their Lagrangian form. The Galerkin approach (5) and the collocation approach (6) result in modified general Nyström methods. Conditions for the initial values guarantee a locally unique numerical solution and the convergence of the method, see Section 4. Finally, in Section 5 we transfer additional geometrical structures to the numerical solution. Criteria for preserving linear and angular momentum, symmetry and reversibility are given.

The results that have been obtained in Section 4 for the energy conserving time integration of conservative systems may in principle be extended to methods with prescribed energy decay for dissipative systems [3]. In comparison with classical fixed step size symplectic integrators [8, Chapters VI and IX], a potential drawback of the proposed generalised Nyström methods is the loss of symplecticity for $\kappa \neq 1$. However, variable step size implementations and the extension to energy decaying schemes give additional flexibility that makes the approach attractive for technical simulations in industrial applications.

6. Simulation results

In this section we present numerical results of the modified Nyström methods applied to the outer solar system described in Section 2. For all investigated Nyström methods we use the three variants “standard”, “projected” and “modified”. The “standard” variant is the original Nyström method, the “projected”

method denotes the projection onto the manifold of constant energy, see Section 2. The “modified” variant stands for the modified Nyström method (7) with the energy equation (8).

All simulations are performed using an experimental MATLAB code that is tailored to nonstiff problems. The numerical solutions of the “standard” and “projected” methods are calculated by fixed-point iterations (for the implicit methods) except the projection step itself that is solved by simplified Newton iterations. The solution of the implicit system (7a), (8) is obtained by a staggered algorithm with simplified Newton iterations as outer loop for determination of κ from Eq. (8). The inner loop computes Υ from Eq. (7a) for a given κ by fixed-point iterations (for the implicit methods).

The outer solar system is a nontrivial nonstiff example of moderate size and was selected because an efficient implementation of the new methods for larger systems is not yet available. The numerical tests with our experimental code show, however, the favourable *qualitative* behaviour of the methods in comparison to the well-known Runge–Kutta/Nyström methods and projection methods.

In Fig. 2 some results for the Galerkin and collocation approach with Radau IIA-quadrature illustrate the superconvergence Theorem 4.5 and the convergence for methods with the energy equation (Theorem 4.6). For the symplectic standard Lobatto IIIA–IIIB pair and the symplectic Gauss methods we get the results of Fig. 3. It is clearly visible, that the global error is neither reduced by the modification nor by the projection. For the Radau IIA methods the projected variant has the smallest error in the solution for this example.

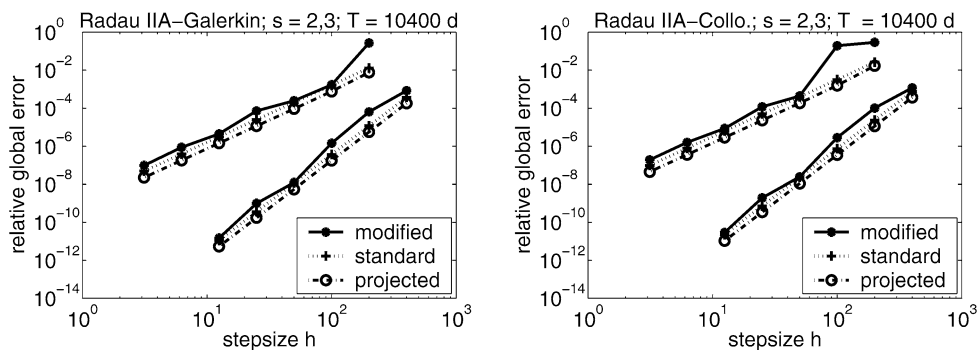


Fig. 2. Outer solar system integrated for 10 400 days by Galerkin and collocation approach with Radau IIA-quadrature.

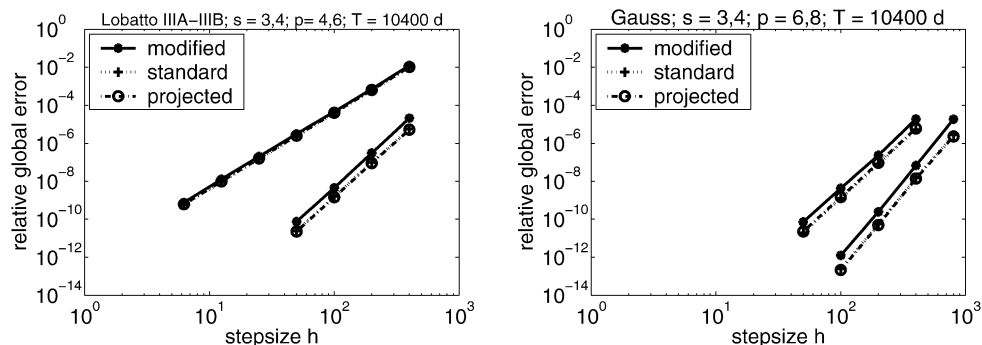


Fig. 3. Outer solar system integrated for 10 400 days by Lobatto and Gauss methods.

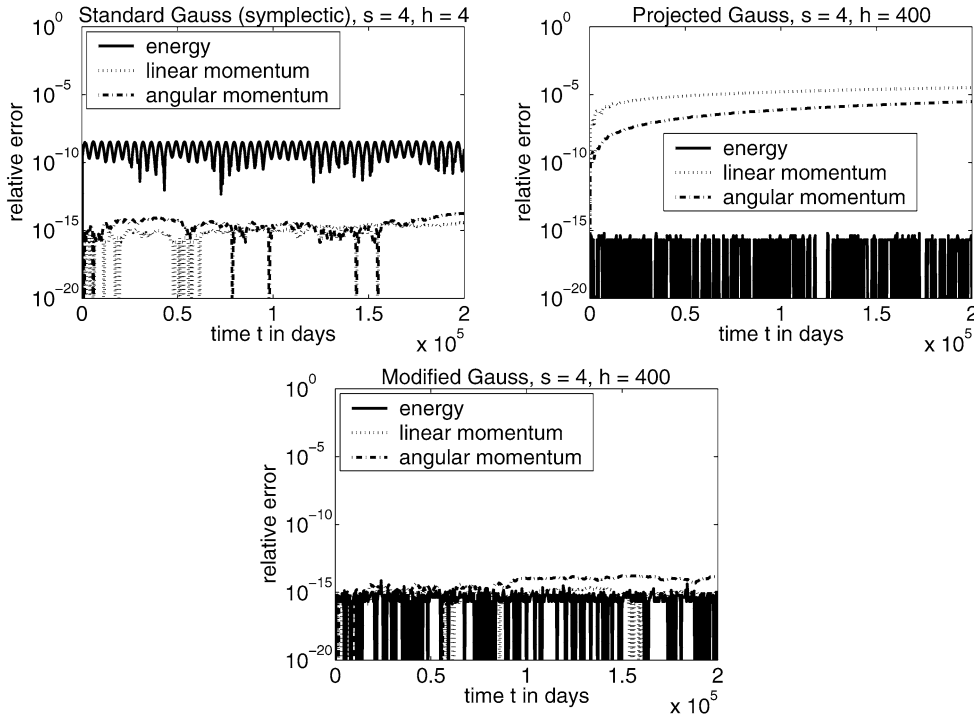


Fig. 4. Outer solar system integrated from 0 to 200 000 days with Gauss method and $s = 4$ stages.

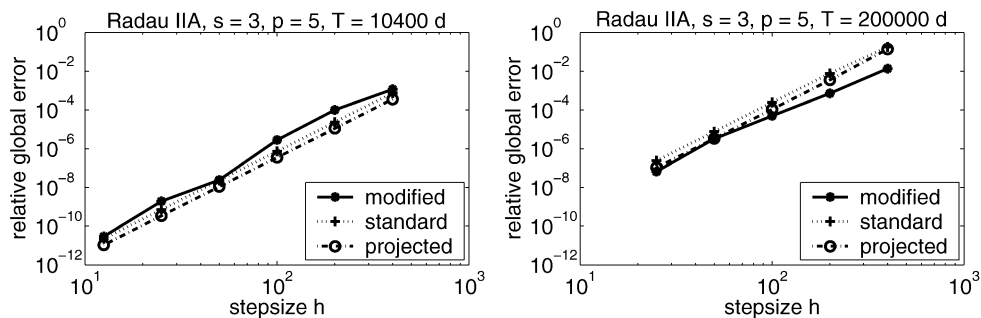


Fig. 5. Outer solar system integrated with Radau IIA collocation methods.

Now we consider the conservation of the invariants for the Gauss method with $s = 4$ stages in more detail. The coefficients of the Gauss methods fulfil the conditions in Theorem 5.1 as can be seen also from the results in Fig. 4. The standard Gauss method preserves linear and angular momentum whereas the energy is not conserved but the energy error is bounded, which is typical for symplectic integrators [8, Chapter X]. Projection onto constant energy causes the effect mentioned in Section 2: linear and angular momentum are not longer preserved. The modified Gauss method combines the positive properties of the standard and the projected method, shown in the lower plot of Fig. 4.

In Figs. 2 and 3 the errors of the modified integration methods are often larger than the errors of the standard methods. But these results are only characteristic for short simulation horizons. In comparison

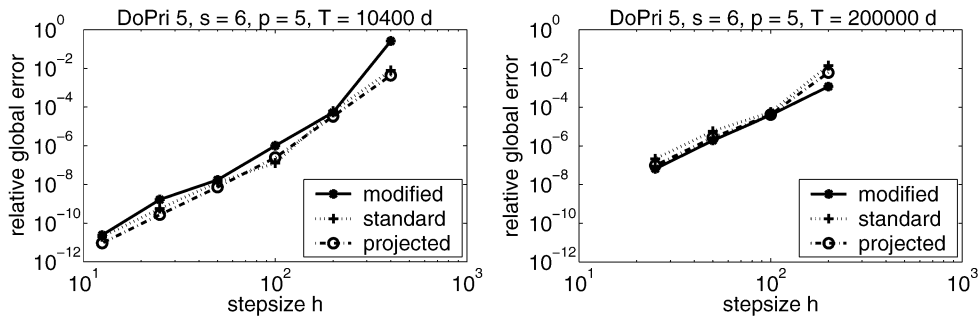


Fig. 6. Outer solar system integrated with method of Dormand and Prince.

we look at the simulation with the Radau IIA method (Fig. 5) and the explicit method from Dormand–Prince of order 5 (Fig. 6). We tested short (10 400 days) and long (200 000 days) simulation times and came to the conclusion, that the errors of the modified variants—especially for large step sizes—are the smallest ones. Just for large step sizes the modified energy conserving methods are interesting, because for standard methods the energy error converges like $\mathcal{O}(h^p)$ anyway, if h tends to 0.

7. Summary

It is well known, that geometric time integration is suitable to solve Hamiltonian systems numerically [8]. Considering the analytical properties of a Hamiltonian system, we try to construct numerical integration methods for Hamiltonian systems in their Lagrangian form. As a starting point we demand the preservation of the total energy (the Hamiltonian) of the numerical solution. A modified Galerkin approach yields modified Nyström methods. In addition a modified collocation approach results in the generalised Nyström form.

We prove the existence of a unique numerical solution and the preservation of linear and quadratic invariants (linear and angular momentum) without projection techniques. Important are the convergence results, which show that the modified methods converge without order reduction. Superconvergence for the Galerkin approach and quadrature formulas with high order is proved. The flow properties symmetry and reversibility can be transferred from the standard Nyström methods to the modified ones.

Finally, numerical experiments with a nonstiff problem and several different methods validate the theoretical statements regarding structure preservation and convergence. For symplectic methods the modified variants do not achieve more accurate results, but long-term simulations point out that certain classical Nyström methods may have large errors that are reduced substantially using the modified methods.

References

- [1] M. Arnold, A. Pfeiffer, Energy conservation in modified Nyström methods for separable Hamiltonian systems, *Proc. Angew. Math. Mech.* 3 (2003) 557–558.
- [2] V.I. Arnol'd, *Mathematische Methoden der klassischen Mechanik*, Birkhäuser, Basel, 1988.
- [3] O.A. Bauchau, C.L. Bottasso, On the design of energy preserving and decaying schemes for flexible, nonlinear multi-body systems, *Comput. Methods Appl. Mech. Engrg.* 169 (1999) 61–79.

- [4] P. Betsch, P. Steinmann, Inherently energy conserving time finite elements for classical mechanics, *J. Comput. Phys.* 160 (2000) 88–116.
- [5] P. Betsch, P. Steinmann, Conservation properties of a time FE method. Part I: Time-stepping schemes for N -body problems, *Internat. J. Numer. Meth. Engrg.* 49 (2000) 599–638.
- [6] P. Betsch, P. Steinmann, Conservation properties of a time FE method. Part II: Time-stepping schemes for non-linear elastodynamics, *Internat. J. Numer. Meth. Engrg.* 50 (2001) 1931–1955.
- [7] P. Deuffhard, A. Hohmann, *Numerical Analysis in Modern Scientific Computing*, second ed., Springer, New York, 2003.
- [8] E. Hairer, Ch. Lubich, G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer, Berlin, 2002.
- [9] E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, second ed., Springer, Berlin, 1993.
- [10] A. Pfeiffer, *Modifizierte Nyström-Verfahren zur geometrischen Zeitintegration von Hamiltonsystemen*, Diplomarbeit, Munich University of Technology, Centre for Mathematical Sciences, 2003.